Variabile casuale Normale



Carl F. Gauss

La var. casuale Normale (o *Gaussiana*) è considerata la più importante "distribuzione" Statistica per le innumerevoli Applicazioni e per le rilevanti proprietà di cui gode

L'importanza di tale v.c. risiede negli indubbi vantaggi formali, ma anche nel fatto che moltissimi fenomeni empirici possono essere rappresentati con un modello di tipo gaussiano

funzione di densità

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$-\infty<x<+\infty$

$e=2,71$

$\pi=3,14$

È una v.c. continua che può assumere valori su tutto l'asse reale

Si dimostra che la f(X) è una funzione di densità perché

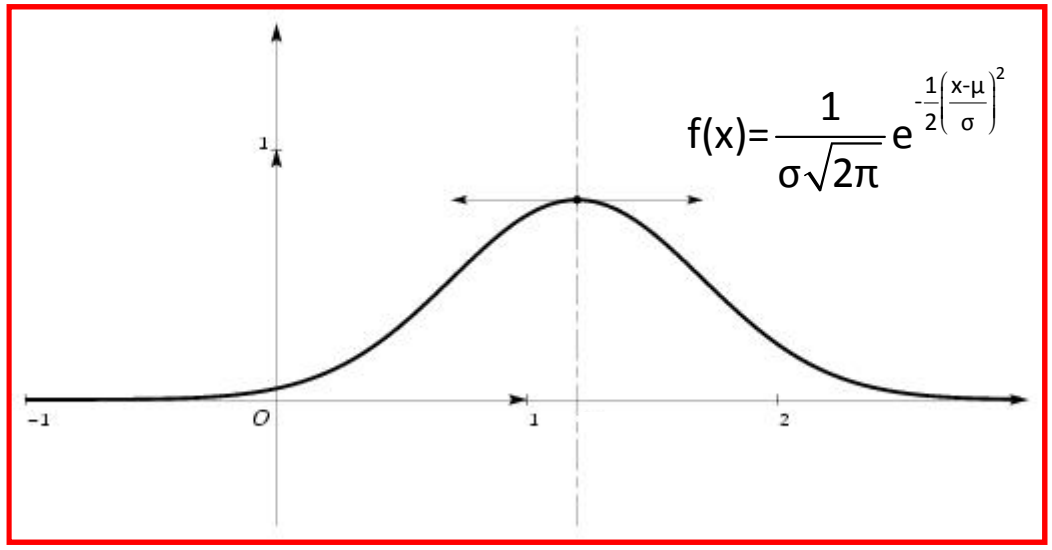
① $f(x)\geq 0$ ② $\int_{-\infty}^{+\infty} f(X)\,dx=1$

$$X \sim N(\mu, \sigma^2)$$

IMPORTANZA DELLA NORMALE

- DIVERSI FENOMENI CONTINUI SEMBRANO SEGUIRE, ALMENO APPROSSIMATIVAMENTE UNA DISTRIBUZIONE NORMALE.
- Può ESSERE UTILIZZATA PER APPROSSIMARE NUMEROSE DISTRIBUZIONI DI PROBABILITà DISCRETE.
- E' ALLA BASE DELL'**INFERENZA STATISTICA CLASSICA** IN VIRTU' DEL **TEOREMA DEL LIMITE CENTRALE**

Forma e parametri della distribuzione



① La funzione di densità della v.c. è simmetrica rispetto al centro

② La densità diminuisce man mano che ci si allontana dal centro (asintoticamente a destra e a sinistra)

L'area sottesa alla curva, probabilità che la var. casuale X assuma valori nel suo dominio, è pari a 1

Si dimostra che i parametri caratteristici della v.c. Normale sono esattamente il valore atteso e la varianza

$$E(X)=\mu \qquad -\infty<\mu<+\infty$$

$$V(X)=\sigma^2 \qquad \sigma^2>0$$

In corrispondenza di $\mathbf{x = \mu \pm \sigma}$ la funzione di prob. Presenta due punti di flesso

PROPRIETA' DELLA NORMALE

- HA UNA FORMA CAMPANULARE E SIMMETRICA

- LE SUE MISURE DI POSIZIONE CENTRALE (VALORE ATTESO, MEDIANA, MODA, MIDRANGE, MEDIA INTERQUARTILE) COINCIDONO.

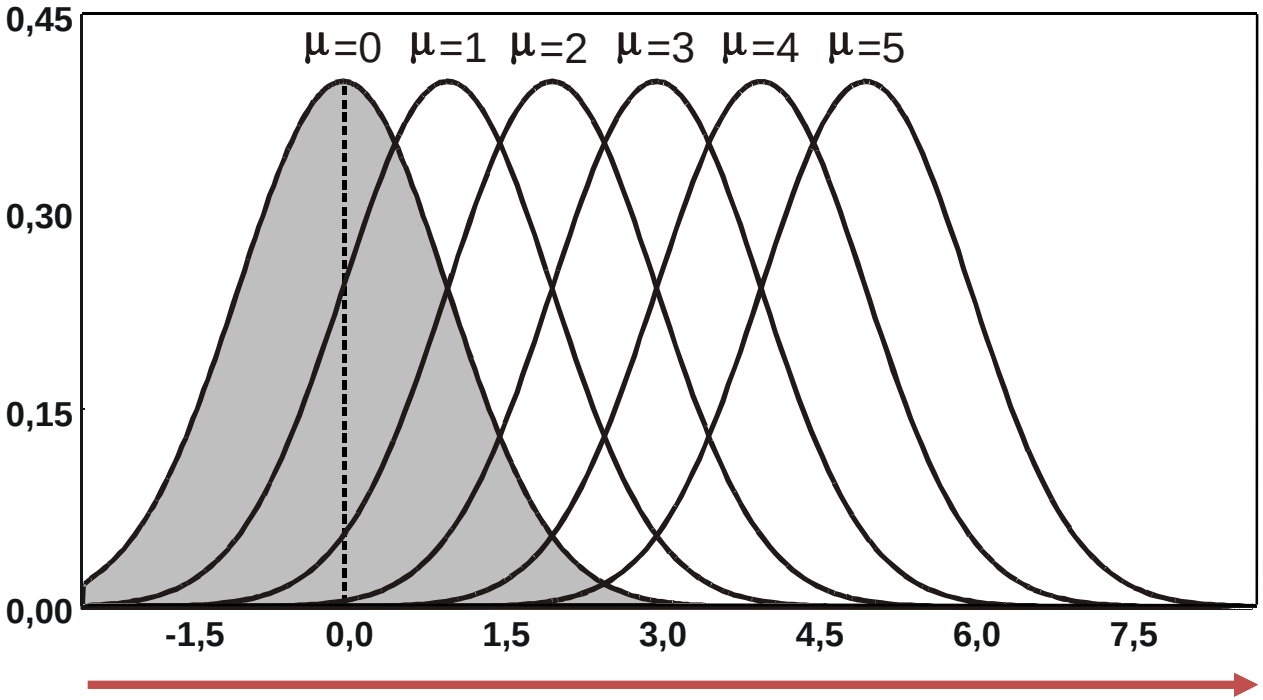
- IL SUO RANGE INTERQUARTILE è PARI A 1.33 VOLTE LO SCARTO QUADRATICO MEDIO, CIOè COPRE UN INTERVALLO COMPRESO TRA :

$$\mu - 2/3\sigma \ , \mu + 2/3\sigma$$

- LA V. ALEATORIA CON DISTRIBUZIONE NORMALE ASSUME VALORI SU TUTTO L'ASSE REALE

Il parametro μ (valore atteso)

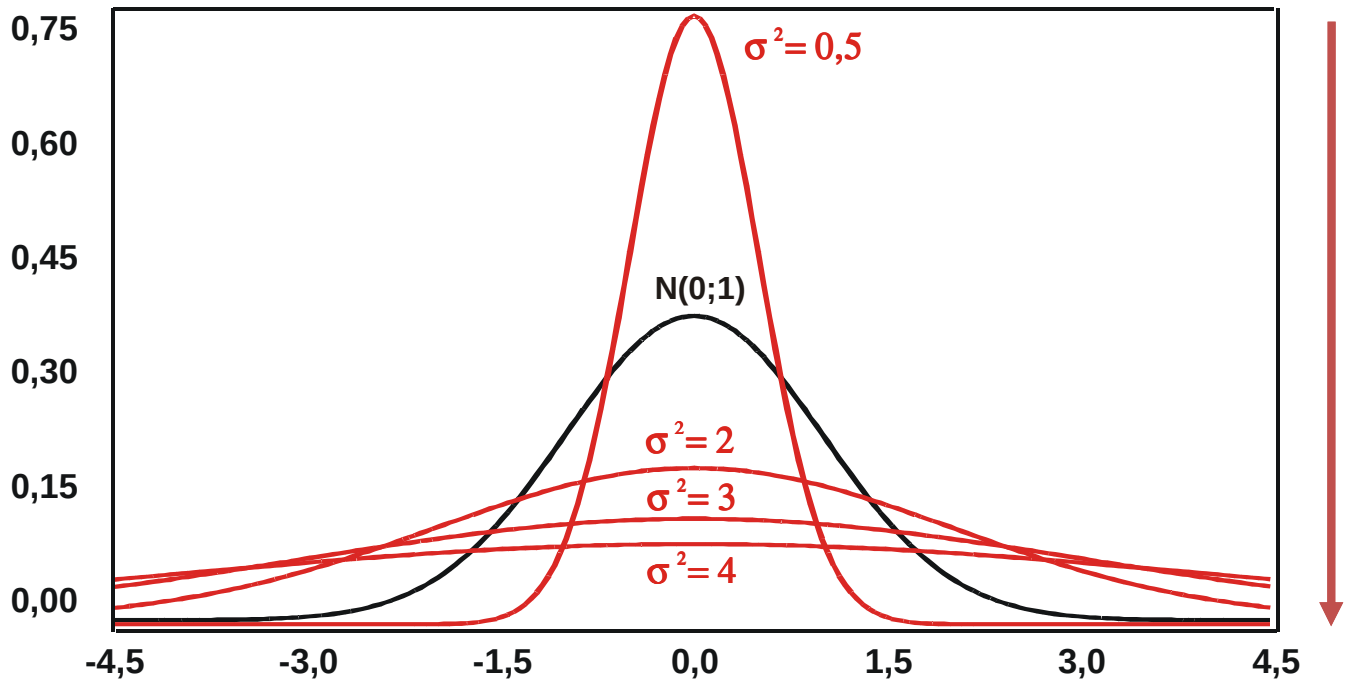
Al variare di μ il grafico resta inalterato nella sua forma ma si modifica solo la sua localizzazione: al crescere di μ la funzione di probabilità si sposta a destra, al diminuire di μ la funzione di probabilità si sposta a sinistra



La distribuzione normale è simmetrica rispetto al suo centro (*valore atteso*): a tale valore centrale corrisponde anche quello più probabile (*valore modale*)

Il parametro σ (deviazione standard)

Al variare di σ il grafico resta inalterato nella sua localizzazione ma si modifica nella sua forma: al crescere di σ la funzione di prob. è più schiacciata rispetto al centro, al diminuire di σ la funzione di prob. è più appuntita



La spiegazione risiede nel fatto che l'area sottesa alla funzione f(x) è sempre pari a 1: se la deviazione standard è bassa i valori di X si addensano intorno al valore atteso e ciò produce un innalzamento dell'ordinata del valore modale

La funzione di ripartizione

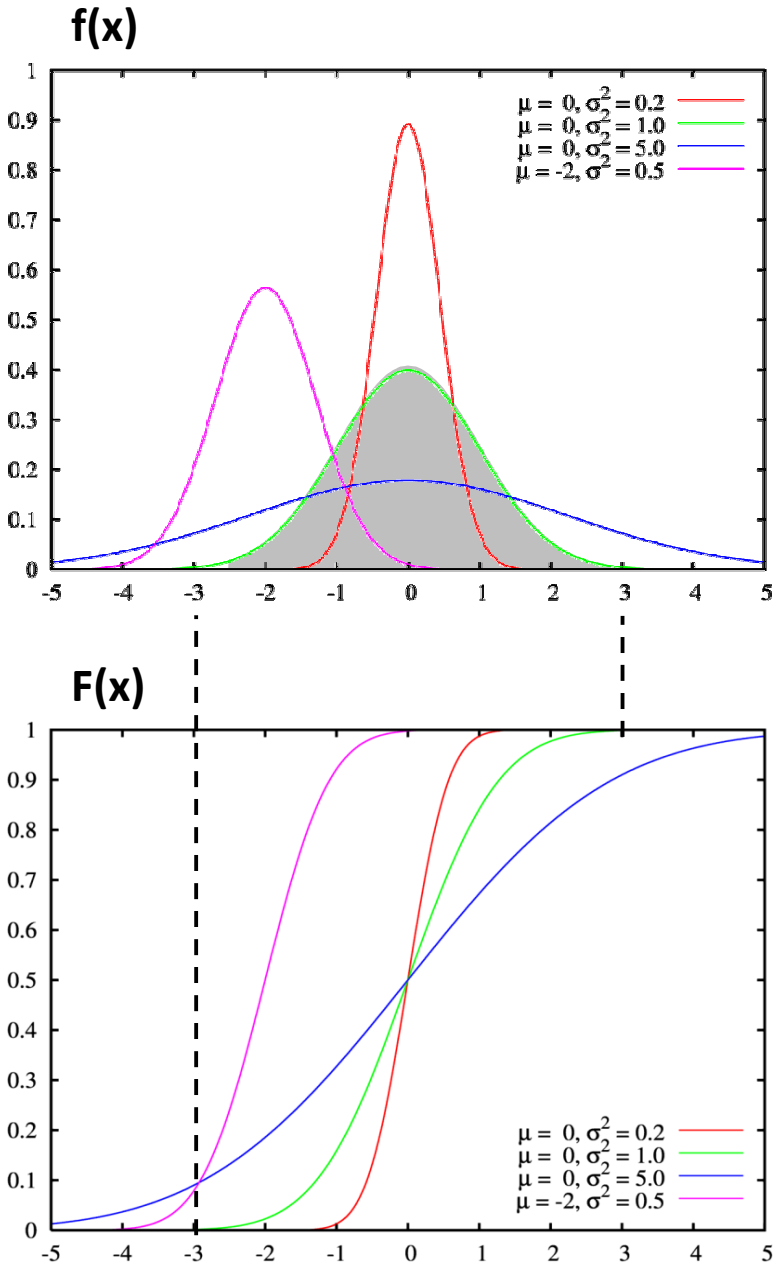
La funzione di ripartizione della v.c. Normale è data formalmente da

$$F(x)=\int_{-\infty}^{x}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}dt$$

Non è facile risolvere questo integrale, tanto che la funzione normale è citata in letteratura come esempio di *funzione non elementare*

Possiamo però descrivere completamente la distribuzione di X attraverso i due parametri μ e σ^2: noti questi valori è possibile calcolare la corrispondente F(x)

Per semplificare il calcolo è possibile far uso di tabelle particolari riconducendo i valori X ad una forma standard



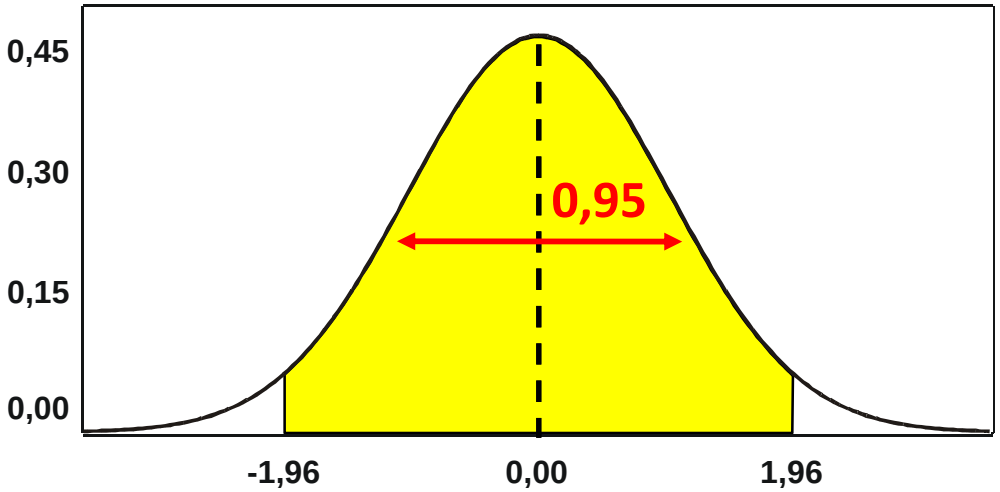
La variabile normale standardizzata

Se la v.c. X ha una distribuzione normale con parametri μ e σ^2, allora la v.c. Z

$$Z = \frac{X - \mu}{\sigma}$$

è ancora una v.c. Normale con media nulla e varianza unitaria

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$



$$E(Z) = 0$$

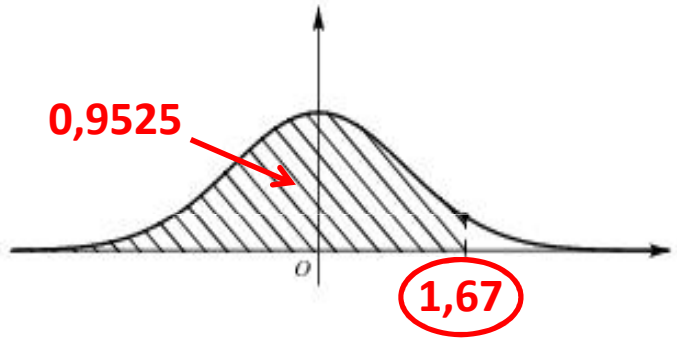
$$V(Z) = 1$$

$$Z \sim N(1,0)$$

Funzione di ripartizione della standardizzata

Permette di semplificare i calcoli delle aree sottese alla funzione di densità

$$P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = P(Z \le z) = \Phi(z) \qquad \forall\, z \ge 0$$



Nella tavola troviamo i valori di Φ(z), funzione di ripartizione di una Normale standardizzata:

sulle righe della tabella si trova il valore intero e la prima cifra decimale, sulle colonne invece la seconda cifra decimale

Valori della f.r. $\Phi(z) = P(Z \le z)$ della v.c. $Z \sim N(0,1)$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767

Generalizziamo…

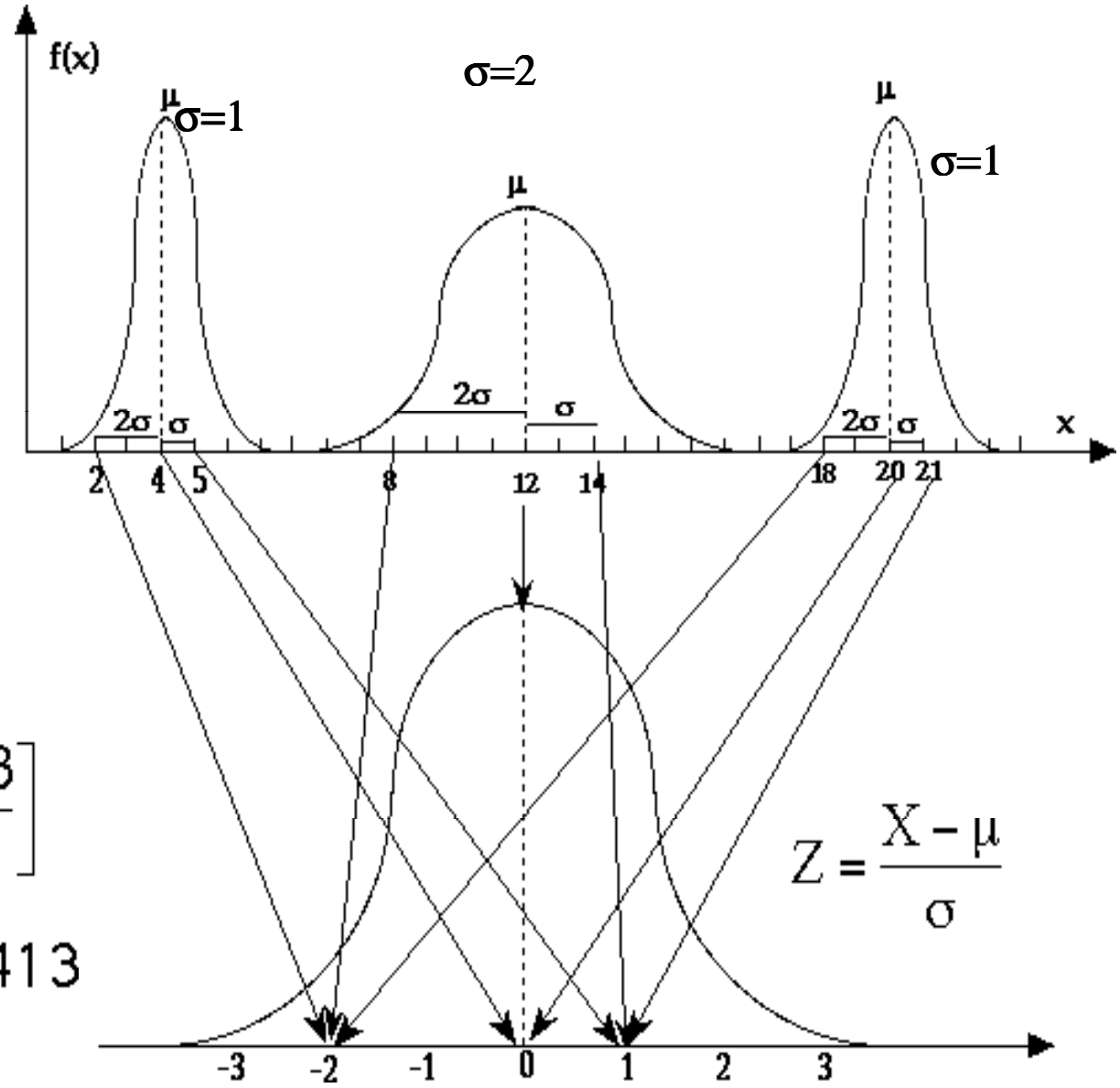
Le tre Normali riportate qui di fianco possono essere di fatto ricondotte a una sola distribuzione, attraverso la trasformazione dei valori x in unità standard

Le aree sottese a $X \sim N(\mu, \sigma^2)$ sono identiche a quelle della $Z \sim N(0,1)$

$$X \sim N(3,4)$$

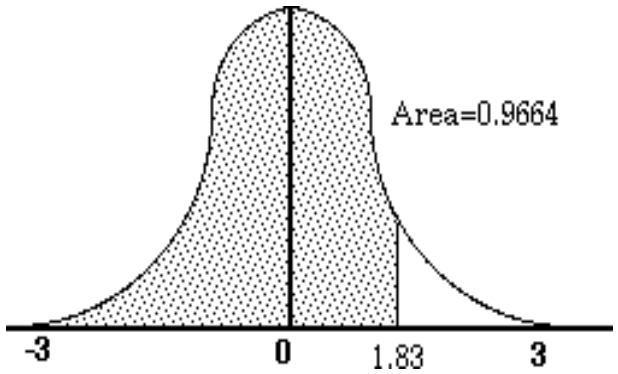
$$P(X < 5) = P\left[\frac{X-3}{2} \leq \frac{5-3}{2}\right]$$

$$= P(Z \leq 1) = 0.8413$$



Esempio

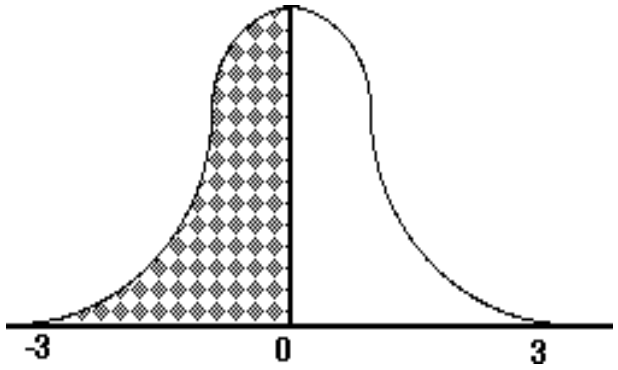
Esempio_1: calcolare l'area compresa fra $-\infty$ e 1.83, in simboli: $\phi(1.83)$



$\phi(1.83) = 0.9664$

$P(Z \leq 1.83)$

Esempio_2: calcolare l'area compresa fra $-\infty$ e 0, in simboli: $\phi(0)$



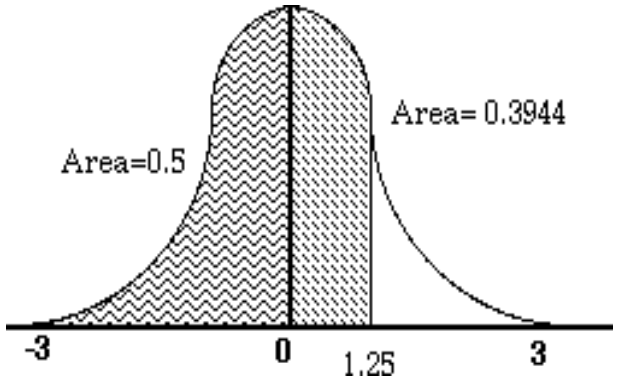
$2\phi(0)=1$ e quindi $\phi(0)=\frac{1}{2}$

$P(Z \leq 0)$

Esempio

Esempio_3: calcolo dell'area tra -∞ e 1.25

$$P(Z \leq 1.25)$$

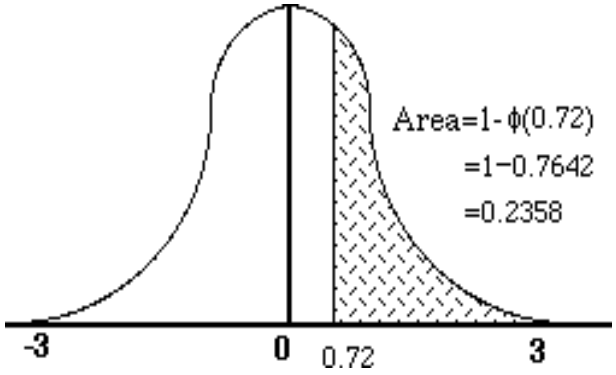


Alcune tavole danno invece l'area compresa tra il valore dato e lo zero. Per ottenere l'intera area occorre sommare 0.5 che è l'area a sinistra dello zero.

$\phi(1.25) = 0.5 + 0.3944 = 0.8944$

Esempio_4: calcolo dell'area a destra di 0.72

$$P(Z > 0.72)$$



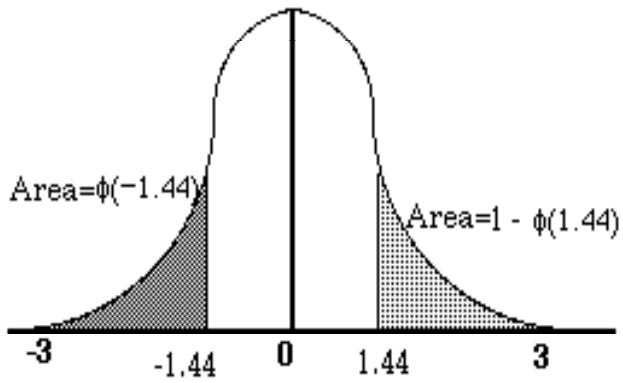
Poiché l'area totale è uguale ad uno si ha l'identità:

Area a sinistra=1- Area a destra

$$1 - P(Z \leq 0.72)$$

Esempio

Esempio_5: calcolo dell'area per valori negativi: $\phi(-1.44)$ $P(Z \leq -1.44)$



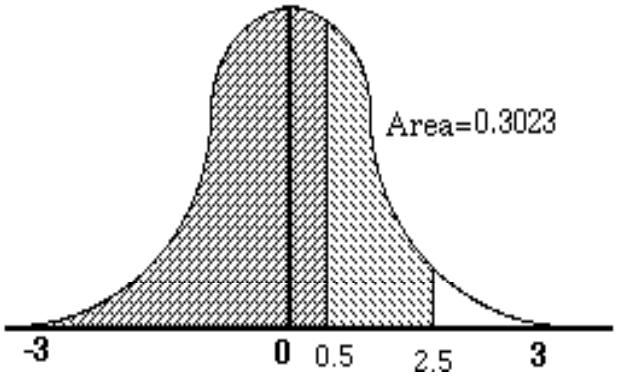
La curva normale è simmetrica per cui

$$\phi(-X) = 1 - \phi(X)$$

Per calcolare l'area a sinistra di un valore negativo si usa l'area a destra del valore positivo corrispondente:

$\phi(-1.44) = 1 - \phi(1.44) = 1 - 0.9251 = 0.0749$

Esempio_6: calcolo dell'area compresa tra due valori positivi. $P(0.5 \leq Z \leq 2.5)$



L'area compresa tra due valori qualsiasi si ottiene per differenza tra le aree a sinistra degli estremi dell'intervallo

Area=$\phi(2.5) - \phi(0.5)$ =0.9938 - 0.6915 =0.3023